

OIP E JCT09
 MAR 18 2003
 TRADEMARK OFFICE

SEQUENCE LISTING

70580

<110> Wood, Kenneth W.
 Sakowicz, Roman
 Goldstein, Lawrence S.B.
 Cleveland, Don W.
 The Regents of the University of California

 <120> Plus End-Directed Microtubule Motor Required for
 Chromosome Congression

 <130> 18557C-000110US

 <140> US 09/150,867
 <141> 1998-09-10

 <150> US 60/058,645
 <151> 1997-09-11

 <160> 11

 <170> PatentIn Ver. 2.0

 <210> 1
 <211> 2954
 <212> PRT
 <213> Xenopus sp.

 <220>
 <223> Xenopus centromere-associated protein-E (XCENP-E)
 member of the kinesin superfamily of microtubule
 motor proteins

 <220>
 <221> DOMAIN
 <222> (1)..(472)
 <223> kinesin like motor domain

 <220>
 <221> DOMAIN
 <222> (473)..(2752)
 <223> rod domain

 <220>
 <221> DOMAIN
 <222> (2753)..(2954)
 <223> tail domain

 <400> 1
 Met Ser Glu Gly Asp Ala Val Lys Val Cys Val Arg Val Arg Pro Leu
 1 5 10 15

 Ile Gln Arg Glu Gln Gly Asp Gln Ala Asn Leu Gln Trp Lys Ala Gly
 20 25 30

 Asn Asn Thr Ile Ser Gln Val Asp Gly Thr Lys Ser Phe Asn Phe Asp
 35 40 45

 Arg Val Phe Asn Ser His Glu Ser Thr Ser Gln Ile Tyr Gln Glu Ile
 50 55 60

58

Ala Val Pro Ile Ile Arg Ser Ala Leu Gln Gly Tyr Asn Gly Thr Ile
65 70 75 80

Phe Ala Tyr Gly Gln Thr Ser Ser Gly Lys Thr Tyr Thr Met Met Gly
85 90 95

Thr Pro Asn Ser Leu Gly Ile Ile Pro Gln Ala Ile Gln Glu Val Phe
100 105 110

Lys Ile Ile Gln Glu Ile Pro Asn Arg Glu Phe Leu Leu Arg Val Ser
115 120 125

Tyr Met Glu Ile Tyr Asn Glu Thr Val Lys Asp Leu Leu Cys Asp Asp
130 135 140

Arg Arg Lys Lys Pro Leu Glu Ile Arg Glu Asp Phe Asn Arg Asn Val
145 150 155 160

Tyr Val Ala Asp Leu Thr Glu Glu Leu Val Met Val Pro Glu His Val
165 170 175

Ile Gln Trp Ile Lys Lys Gly Glu Lys Asn Arg His Tyr Gly Glu Thr
180 185 190

Lys Met Asn Asp His Ser Ser Arg Ser His Thr Ile Phe Arg Met Ile
195 200 205

Val Glu Ser Arg Asp Arg Asn Asp Pro Thr Asn Ser Glu Asn Cys Asp
210 215 220

Gly Ala Val Met Val Ser His Leu Asn Leu Val Asp Leu Ala Gly Ser
225 230 235 240

Glu Arg Ala Ser Gln Thr Gly Ala Glu Gly Val Arg Leu Lys Glu Gly
245 250 255

Cys Asn Ile Asn Arg Ser Leu Phe Ile Leu Gly Gln Val Ile Lys Lys
260 265 270

Leu Ser Asp Gly Gln Ala Gly Gly Phe Ile Asn Tyr Arg Asp Ser Lys
275 280 285

Leu Thr Arg Ile Leu Gln Asn Ser Leu Gly Gly Asn Ala Lys Thr Val
290 295 300

Ile Ile Cys Thr Ile Thr Pro Val Ser Phe Asp Glu Thr Leu Ser Thr
305 310 315 320

Leu Gln Phe Ala Ser Thr Ala Lys His Val Arg Asn Thr Pro His Val
325 330 335

Asn Glu Val Leu Asp Asp Glu Ala Leu Leu Lys Arg Tyr Arg Lys Glu
340 345 350

Ile Leu Asp Leu Lys Lys Gln Leu Glu Asn Leu Glu Ser Ser Ser Glu
355 360 365

Thr Lys Ala Gln Ala Met Ala Lys Glu Glu His Thr Gln Leu Leu Ala
370 375 380

Glu Ile Lys Gln Leu His Lys Glu Arg Glu Asp Arg Ile Trp His Leu
385 390 395 400

Thr Asn Ile Val Val Ala Ser Ser Gln Glu Ser Gln Gln Asp Gln Arg
405 410 415

Val Lys Arg Lys Arg Arg Val Thr Trp Ala Pro Gly Lys Ile Gln Asn
420 425 430

Ser Leu His Ala Ser Gly Val Ser Asp Phe Asp Met Leu Ser Arg Leu
435 440 445

Pro Gly Asn Phe Ser Lys Lys Ala Lys Phe Ser Asp Met Pro Ser Phe
450 455 460

Pro Glu Ile Asp Asp Ser Val Cys Thr Glu Phe Ser Asp Phe Asp Asp
465 470 475 480

Ala Leu Ser Met Met Asp Ser Asn Gly Ile Asp Ala Glu Trp Asn Leu
485 490 495

Ala Ser Lys Val Thr His Arg Glu Lys Thr Ser Leu His Gln Ser Met
500 505 510

Ile Asp Phe Gly Gln Ile Ser Asp Ser Val Gln Phe His Asp Ser Ser
515 520 525

Lys Glu Asn Gln Leu Gln Tyr Leu Pro Lys Asp Ser Gly Asp Met Ala
530 535 540

Glu Cys Arg Lys Ala Ser Phe Glu Lys Glu Ile Thr Ser Leu Gln Gln
545 550 555 560

Gln Leu Gln Ser Lys Glu Glu Glu Lys Lys Glu Leu Val Gln Ser Phe
565 570 575

Glu Leu Lys Ile Ala Glu Leu Glu Glu Gln Leu Ser Val Lys Ala Lys
580 585 590

Asn Leu Glu Met Val Thr Asn Ser Arg Glu His Ser Ile Asn Ala Glu
595 600 605

Val Gln Thr Asp Val Glu Lys Glu Val Val Arg Lys Glu Met Ser Val
610 615 620

Leu Gly Asp Ser Gly Tyr Asn Ala Ser Asn Ser Asp Leu Gln Asp Ser
625 630 635 640

Ser Val Asp Gly Lys Arg Leu Ser Ser Ser His Asp Glu Cys Ile Glu
645 650 655

His Arg Lys Met Leu Glu Gln Lys Ile Val Asp Leu Glu Glu Phe Ile
660 665 670

Glu Asn Leu Asn Lys Lys Ser Glu Asn Asp Lys Gln Lys Ser Ser Glu
675 680 685

Gln Asp Phe Met Glu Ser Ile Gln Leu Cys Glu Ala Ile Met Ala Glu
690 695 700

Lys Ala Asn Ala Leu Glu Glu Leu Ala Leu Met Arg Asp Asn Phe Asp
 705 710 715 720
 Asn Ile Ile Leu Glu Asn Glu Thr Leu Lys Arg Glu Ile Ala Asp Leu
 725 730 735
 Glu Arg Ser Leu Lys Glu Asn Gln Glu Thr Asn Glu Phe Glu Ile Leu
 740 745 750
 Glu Lys Glu Thr Gln Lys Glu His Glu Ala Gln Leu Ile His Glu Ile
 755 760 765
 Gly Ser Leu Lys Lys Leu Val Glu Asn Ala Glu Met Tyr Asn Gln Asn
 770 775 780
 Leu Glu Glu Asp Leu Glu Thr Lys Thr Lys Leu Leu Lys Glu Gln Glu
 785 790 795 800
 Ile Gln Leu Ala Glu Leu Arg Lys Arg Ala Asp Asn Leu Gln Lys Lys
 805 810 815
 Val Arg Asn Phe Asp Leu Ser Val Ser Met Gly Asp Ser Glu Lys Leu
 820 825 830
 Cys Glu Glu Ile Phe Gln Leu Lys Gln Ser Leu Ser Asp Ala Glu Ala
 835 840 845
 Val Thr Arg Asp Ala Gln Lys Glu Cys Ser Phe Leu Arg Ser Glu Asn
 850 855 860
 Leu Glu Leu Lys Glu Lys Met Glu Asp Thr Ser Asn Trp Tyr Asn Gln
 865 870 875 880
 Lys Glu Lys Ala Ala Ser Leu Phe Glu Lys Gln Leu Glu Thr Glu Lys
 885 890 895
 Ser Asn Tyr Lys Lys Met Glu Ala Asp Leu Gln Lys Glu Leu Gln Ser
 900 905 910
 Ala Phe Asn Glu Ile Asn Tyr Leu Asn Gly Leu Leu Ala Gly Lys Val
 915 920 925
 Pro Arg Asp Leu Leu Ser Arg Val Glu Leu Glu Lys Lys Val Ser Glu
 930 935 940
 Phe Ser Lys Gln Leu Glu Lys Ala Leu Glu Glu Lys Asn Ala Leu Glu
 945 950 955 960
 Asn Glu Val Thr Cys Leu Ser Glu Tyr Lys Phe Leu Pro Asn Glu Val
 965 970 975
 Glu Cys Leu Lys Asn Gln Ile Ser Lys Ala Ser Glu Glu Ile Met Leu
 980 985 990
 Leu Lys Gln Glu Gly Glu His Ser Ala Ser Ile Ile Ser Lys Gln Glu
 995 1000 1005
 Ile Ile Met Gln Glu Gln Ser Glu Gln Ile Leu Gln Leu Thr Asp Glu
 1010 1015 1020

Val Thr His Thr Gln Ser Lys Val Gln Gln Thr Glu Glu Gln Tyr Leu
 1025 1030 1035 1040
 Glu Met Lys Lys Met His Asp Asp Leu Phe Glu Lys Tyr Ile Arg Asn
 1045 1050 1055
 Lys Ser Glu Ala Glu Asp Leu Leu Arg Glu Met Glu Asn Leu Lys Gly
 1060 1065 1070
 Thr Met Glu Ser Val Glu Val Lys Ile Ala Asp Thr Lys His Glu Leu
 1075 1080 1085
 Glu Glu Thr Ile Arg Asp Lys Glu Gln Leu Leu His Glu Lys Lys Tyr
 1090 1095 1100
 Phe Phe Gln Ala Met Gln Thr Ile Phe Pro Ile Thr Pro Leu Ser Asp
 1105 1110 1115 1120
 Ser Leu Pro Pro Ser Lys Leu Val Glu Gly Asn Ser Gln Asp Pro Ile
 1125 1130 1135
 Glu Ile Asn Asp Tyr His Asn Leu Ile Ala Leu Ala Thr Glu Arg Asn
 1140 1145 1150
 Asn Ile Met Val Cys Leu Glu Thr Glu Arg Asn Ser Leu Lys Glu Gln
 1155 1160 1165
 Val Ile Asp Leu Asn Thr Gln Leu Gln Ser Leu Gln Ala Gln Ser Ile
 1170 1175 1180
 Glu Lys Ser Asp Leu Gln Lys Pro Lys Gln Asp Leu Glu Glu Gly Glu
 1185 1190 1195 1200
 Val Lys Leu Leu Leu Glu Met Glu Leu Leu Lys Gly His Leu Thr Asp
 1205 1210 1215
 Ser Gln Leu Ser Ile Glu Lys Leu Gln Leu Glu Asn Leu Glu Val Thr
 1220 1225 1230
 Glu Lys Leu Gln Thr Leu Gln Glu Glu Met Lys Asn Ile Thr Ile Glu
 1235 1240 1245
 Arg Asn Glu Leu Gln Thr Asn Phe Glu Asp Leu Lys Ala Glu His Asp
 1250 1255 1260
 Ser Leu Lys Gln Asp Leu Ser Glu Asn Ile Glu Gln Ser Ile Glu Thr
 1265 1270 1275 1280
 Gln Asp Glu Leu Arg Ala Ala Gln Glu Glu Leu Arg Glu Gln Lys Gln
 1285 1290 1295
 Leu Val Asp Ser Phe Arg Gln Gln Leu Leu Asp Cys Ser Val Gly Ile
 1300 1305 1310
 Ser Ser Pro Asn His Asp Ala Val Ala Asn Gln Glu Lys Val Ser Leu
 1315 1320 1325
 Gly Glu Val Asn Ser Leu Gln Ser Glu Met Leu Arg Gly Glu Arg Asp
 1330 1335 1340

62

Glu Leu Gln Thr Ser Cys Lys Ala Leu Val Ser Glu Leu Glu Leu Leu
 1345 1350 1355 1360

Arg Ala His Val Lys Ser Val Glu Gly Glu Asn Leu Glu Ile Thr Lys
 1365 1370 1375

Lys Leu Asn Gly Leu Glu Lys Glu Ile Leu Gly Lys Ser Glu Glu Ser
 1380 1385 1390

Glu Val Leu Lys Ser Met Leu Glu Asn Leu Lys Glu Asp Asn Asn Lys
 1395 1400 1405

Leu Lys Glu Gln Ala Glu Glu Tyr Ser Ser Lys Glu Asn Gln Phe Ser
 1410 1415 1420

Leu Glu Glu Val Phe Ser Gly Ser Gln Lys Leu Val Asp Glu Ile Glu
 1425 1430 1435 1440

Val Leu Lys Ala Gln Leu Lys Ala Ala Glu Glu Arg Leu Glu Ile Lys
 1445 1450 1455

Asp Arg Asp Tyr Phe Glu Leu Val Gln Thr Ala Asn Thr Asn Leu Val
 1460 1465 1470

Glu Gly Lys Leu Glu Thr Pro Leu Gln Ala Asp His Glu Glu Asp Ser
 1475 1480 1485

Ile Asp Arg Arg Ser Glu Glu Met Glu Ile Lys Val Leu Gly Glu Lys
 1490 1495 1500

Leu Glu Arg Asn Gln Tyr Leu Leu Glu Arg Leu Gln Glu Glu Lys Leu
 1505 1510 1515 1520

Glu Leu Ser Asn Lys Leu Glu Ile Leu Gln Lys Glu Met Glu Thr Ser
 1525 1530 1535

Val Leu Leu Lys Asp Asp Leu Gln Gln Lys Leu Glu Ser Leu Leu Ser
 1540 1545 1550

Glu Asn Ile Ile Leu Lys Glu Asn Ile Asp Thr Thr Leu Lys His His
 1555 1560 1565

Ser Asp Thr Gln Ala Gln Leu Gln Lys Thr Gln Gln Glu Leu Gln Leu
 1570 1575 1580

Ala Lys Asn Leu Ala Ile Ala Ala Ser Asp Asn Cys Pro Ile Thr Gln
 1585 1590 1595 1600

Glu Lys Glu Thr Ser Ala Asp Cys Val His Pro Leu Glu Glu Lys Ile
 1605 1610 1615

Leu Leu Leu Thr Glu Glu Leu His Gln Lys Thr Asn Glu Gln Glu Lys
 1620 1625 1630

Leu Leu His Glu Lys Asn Glu Leu Glu Gln Ala Gln Val Glu Leu Lys
 1635 1640 1645

Cys Glu Val Glu His Leu Met Lys Ser Met Ile Glu Ser Lys Ser Ser
 1650 1655 1660

63

Leu Glu Ser Leu Gln His Glu Lys His Asp Thr Glu Gln Gln Leu Leu
 1665 1670 1675 1680

Ala Leu Lys Gln Gln Met Gln Val Val Thr Gln Glu Lys Lys Glu Leu
 1685 1690 1695

Gln Gln Thr His Glu His Leu Thr Ala Glu Val Asp His Leu Lys Glu
 1700 1705 1710

Asn Ile Glu Leu Gly Leu Asn Phe Lys Asn Glu Ala Gln Gln Lys Thr
 1715 1720 1725

Thr Lys Glu Gln Cys Leu Leu Asn Glu Asn Lys Glu Leu Glu Gln Ser
 1730 1735 1740

Gln His Arg Leu Gln Cys Glu Ile Glu Glu Leu Met Lys Ser Leu Lys
 1745 1750 1755 1760

Asp Lys Glu Ser Ala Leu Glu Thr Leu Lys Glu Ser Glu Gln Lys Val
 1765 1770 1775

Ile Asn Leu Asn Gln Glu Met Glu Met Val Met Leu Glu Met Glu Glu
 1780 1785 1790

Leu Lys Asn Ser Gln Arg Thr Val Ile Ala Glu Arg Asp Gln Leu Gln
 1795 1800 1805

Asp Asp Leu Arg Glu Ser Val Glu Met Ser Ile Glu Thr Gln Asp Asp
 1810 1815 1820

Leu Arg Lys Ala Gln Glu Ala Leu Gln Gln Gln Lys Asp Lys Val Gln
 1825 1830 1835 1840

Glu Leu Thr Ser Gln Ile Ser Val Leu Gln Glu Lys Ile Ser Leu Leu
 1845 1850 1855

Glu Asn Gln Met Leu Tyr Asn Val Ala Thr Val Lys Glu Thr Leu Ser
 1860 1865 1870

Glu Arg Asp Asp Leu Asn Gln Ser Lys Gln His Leu Phe Ser Glu Ile
 1875 1880 1885

Glu Thr Leu Ser Leu Ser Leu Lys Glu Lys Glu Phe Ala Leu Glu Gln
 1890 1895 1900

Ala Glu Lys Asp Lys Ala Asp Ala Ala Arg Lys Thr Ile Asp Ile Thr
 1905 1910 1915 1920

Glu Lys Ile Ser Asn Ile Glu Glu Gln Leu Leu Gln Gln Ala Thr Asn
 1925 1930 1935

Leu Lys Glu Thr Leu Tyr Glu Arg Glu Ser Leu Ile Gln Cys Lys Glu
 1940 1945 1950

Gln Leu Ala Leu Asn Thr Glu His Leu Arg Glu Thr Leu Lys Ser Lys
 1955 1960 1965

Asp Leu Ala Leu Gly Lys Met Glu Gln Glu Arg Asp Glu Ala Ala Asn
 1970 1975 1980

64

Lys Val Ile Ala Leu Thr Glu Lys Met Ser Ser Leu Glu Glu Gln Ile
1985 1990 1995 2000

Asn Glu Asn Val Thr Thr Leu Lys Glu Gly Glu Gly Glu Lys Glu Thr
2005 2010 2015

Phe Tyr Leu Gln Arg Pro Ser Lys Gln Gln Ser Ser Ser Gln Met Glu
2020 2025 2030

Glu Leu Arg Glu Ser Leu Lys Thr Lys Asp Leu Gln Leu Glu Glu Ala
2035 2040 2045

Glu Lys Glu Ile Ser Glu Ala Thr Asn Glu Ile Lys Asn Leu Thr Ala
2050 2055 2060

Lys Ile Ser Ser Leu Glu Glu Glu Ile Leu Gln Asn Ala Ser Ile Leu
2065 2070 2075 2080

Asn Glu Ala Val Ser Glu Arg Glu Asn Leu Arg His Ser Lys Gln Gln
2085 2090 2095

Leu Val Ser Glu Leu Glu Gln Leu Ser Leu Thr Leu Lys Ser Arg Asp
2100 2105 2110

His Ala Phe Ala Gln Ser Lys Arg Glu Lys Asp Glu Ala Val Asn Lys
2115 2120 2125

Ile Ala Ser Leu Ala Glu Glu Ile Lys Ile Leu Thr Lys Glu Met Asp
2130 2135 2140

Glu Phe Arg Asp Ser Lys Glu Ser Leu Gln Glu Gln Ser Ser His Leu
2145 2150 2155 2160

Ser Glu Glu Leu Cys Thr Tyr Lys Thr Glu Leu Gln Met Leu Lys Gln
2165 2170 2175

Gln Lys Glu Asp Ile Asn Asn Lys Leu Ala Glu Lys Val Lys Glu Val
2180 2185 2190

Asp Glu Leu Leu Gln His Leu Ser Ser Leu Lys Glu Gln Leu Asp Gln
2195 2200 2205

Ile Gln Met Glu Leu Arg Asn Glu Lys Leu Arg Asn Tyr Glu Leu Cys
2210 2215 2220

Glu Lys Met Asp Ile Met Glu Lys Glu Ile Ser Val Leu Arg Leu Met
2225 2230 2235 2240

Gln Asn Glu Pro Gln Gln Glu Glu Asp Asp Val Ala Glu Arg Met Asp
2245 2250 2255

Ile Leu Glu Ser Arg Asn Gln Glu Ile Gln Glu Leu Met Glu Lys Ile
2260 2265 2270

Ser Ala Val Tyr Ser Glu Gln His Thr Leu Leu Ser Ser Leu Ser Ser
2275 2280 2285

Glu Leu Gln Lys Glu Thr Glu Ala His Lys His Cys Met Leu Asn Ile
2290 2295 2300

65

Lys Glu Ser Leu Ser Ser Thr Leu Ser Arg Ser Phe Gly Ser Leu Gln
 2305 2310 2315 2320
 Thr Glu His Val Lys Leu Asn Thr Gln Leu Gln Thr Leu Leu Asn Lys
 2325 2330 2335
 Phe Lys Val Val Tyr Arg Thr Ala Ala Val Lys Glu Asp His Ser Leu
 2340 2345 2350
 Ile Lys Asp Tyr Glu Lys Asp Leu Ala Ala Glu Gln Lys Arg His Asp
 2355 2360 2365
 Glu Leu Arg Leu Gln Leu Gln Cys Leu Glu Gln His Gly Arg Lys Trp
 2370 2375 2380
 Ser Asp Ser Ala Ser Glu Glu Leu Lys Phe Cys Glu Ile Glu Phe Leu
 2385 2390 2395 2400
 Asn Glu Leu Leu Phe Lys Lys Ala Asn Ile Ile Gln Ser Val Gln Asp
 2405 2410 2415
 Asp Phe Ser Glu Val Gln Val Phe Leu Asn Gln Val Gly Ser Thr Leu
 2420 2425 2430
 Gln Glu Glu Leu Glu His Lys Lys Gly Phe Met Gln Trp Leu Glu Glu
 2435 2440 2445
 Phe Gly Asp Leu His Val Asp Ala Lys Lys Leu Ser Glu Gly Met Gln
 2450 2455 2460
 Gln Glu Asn Arg Arg Ile Ala Ser Thr Ile Gln Leu Leu Thr Lys Arg
 2465 2470 2475 2480
 Leu Lys Ala Val Val Gln Ser Lys Ile Gln Arg Glu Ile Thr Val Tyr
 2485 2490 2495
 Leu Asn Gln Phe Glu Ala Lys Leu Gln Glu Lys Lys Glu Gln Asn Lys
 2500 2505 2510
 Glu Leu Met Arg Arg Met Glu His His Gly Pro Ser Ala Ser Val Met
 2515 2520 2525
 Glu Glu Glu Asn Ala Arg Leu Leu Gly Ile Leu Lys Thr Val Gln Asp
 2530 2535 2540
 Glu Ser Lys Lys Leu Gln Ser Arg Ile Lys Met Leu Glu Asn Glu Leu
 2545 2550 2555 2560
 Asn Leu Val Lys Asp Asp Ala Met His Lys Gly Glu Lys Val Ala Ile
 2565 2570 2575
 Leu Gln Asp Lys Leu Leu Ser Arg Asn Ala Glu Ala Glu Leu Asn Ala
 2580 2585 2590
 Met Gln Val Lys Leu Thr Lys Lys Gln Asp Asn Leu Gln Ala Ala Met
 2595 2600 2605
 Lys Glu Ile Glu Asn Leu Gln Lys Met Val Ala Lys Gly Ala Val Pro
 2610 2615 2620

Tyr Lys Glu Glu Ile Asp Asn Leu Lys Thr Lys Val Val Lys Ile Glu
 2625 2630 2635 2640
 Met Glu Lys Ile Lys Tyr Ser Lys Ala Thr Asp Gln Glu Ile Ala Tyr
 2645 2650 2655
 Leu Lys Ser Cys Leu Glu Asp Lys Glu Glu Gly Leu Arg Arg Leu Lys
 2660 2665 2670
 Glu Glu Leu Arg Arg Ala Gln Ala Asp Asn Asp Thr Thr Val Cys Val
 2675 2680 2685
 Pro Lys Asp Tyr Gln Lys Ala Ser Thr Phe Pro Val Thr Cys Gly Gly
 2690 2695 2700
 Gly Ser Gly Ile Val Gln Ser Thr Ala Met Leu Val Leu Gln Ser Glu
 2705 2710 2715 2720
 Lys Ala Ala Leu Glu Arg Glu Leu Ser His Tyr Lys Lys Lys Tyr His
 2725 2730 2735
 His Leu Ser Arg Thr Met Ser Ser Ser Glu Asp Arg Lys Lys Thr Lys
 2740 2745 2750
 Ala Lys Ser Asp Ala His Ser Ser His Thr Gly Ser Ser His Arg Gly
 2755 2760 2765
 Ser Pro His Lys Thr Glu Thr Tyr Arg His Gly Pro Val Thr Pro Glu
 2770 2775 2780
 Arg Ser Glu Met Pro Ser Leu His Leu Gly Ser Pro Lys Lys Ser Glu
 2785 2790 2795 2800
 Ser Ser Thr Lys Arg Val Val Ser Pro Asn Arg Ser Glu Ile Tyr Ser
 2805 2810 2815
 Gln Leu Val Met Ser Pro Gly Lys Thr Gly Met His Lys His Ile Leu
 2820 2825 2830
 Ser Pro Ser Lys Val Gly Leu His Lys Lys Arg Ala Leu Ser Pro Asn
 2835 2840 2845
 Arg Ser Glu Met Pro Thr Gln His Val Ile Ser Pro Gly Lys Thr Gly
 2850 2855 2860
 Leu His Lys Asn Leu Thr Glu Ser Thr Leu Phe Asp Asn Leu Ser Ser
 2865 2870 2875 2880
 Pro Cys Lys Gln Gln Lys Val Gln Glu Asn Leu Asn Ser Pro Lys Gly
 2885 2890 2895
 Lys Leu Phe Asp Val Lys Ser Lys Ser Met Pro Tyr Cys Pro Ser Gln
 2900 2905 2910
 Phe Phe Asp Asn Ser Lys Leu Gly Asp Phe Ser Glu Leu Asn Thr Ala
 2915 2920 2925
 Glu Ser Asn Asp Lys Ser Gln Ala Glu Asn Trp Trp Tyr Glu Ala Lys
 2930 2935 2940

Lys Glu Thr Ala Pro Glu Cys Lys Thr Ser
2945 2950

<210> 2
<211> 9626
<212> DNA
<213> Xenopus sp.

<220>
<223> XCENP-E nucleotide sequence

<220>
<221> CDS
<222> (143)..(9007)

<400> 2
gaattccgga gtcggatagg ctagtccggc gagggaaatt caaactgggt atagaagaaa 60
cttgaaccgc cgccaaaagg gactaaagt acagagacag ggagcgggtg cggtagcgat 120
ttctccacta atcgggtctca aaatgtccga gggagatgca gttaaagtgt gtgtgagggg 180
tcggccgctt atacagagag aacaagggga tcaagccaac ctgcaatgga aggctggaaa 240
caacaccatt tcccaagttg atgggacaaa gtctttcaat ttcgatcgtg tatttaattc 300
tcacgaatca acaagtcaaa tttaccaaga aatagcagta cctatcatal gatcagcttt 360
gcagggatat aatggcacia tatttgcata cggacagaca tcttcaggca agacgtacac 420
aatgatggga acaccaaatt cattgggcat aatacccaa gccatacagg aagtttttaa 480
aattattcag gagataccga acagagagtt tcttctaaga gtttcttata tggagattta 540
caatgaaact gtgaaagacc tactgtgtga tgacagaaga aagaagccct tggaaattcg 600
cgaggatttt aatagaaacg tgtatgttgc tgacctgact gaagaacttg taatgggtcc 660
tgaacatgta atacagtga tcaaaaaggg tgaaaaaac agacattatg gagagactaa 720
aatgaatgat catagtagtc gttcacatac aatattttag atgattgttg aaagccgaga 780
cagaaatgat cccacaaatt cagagaactg tgatggagct gtcattggtat ctcacttgaa 840
tttggtagat cttgctggca gtgaaagagc aagccaaact ggagctgaag gtgtgagact 900
taaggaaagg accgacatca accgcagctt gtttatcctt ggacagggtta ttaagaagct 960
tagcgacggc caggctgggtg gatttataaa ctacagagac agcaaactca ccagaattct 1020
ccaaaattca ttgggaggaa atgctaaaac ggttataatt tgcacaatta cgccagtttc 1080
ttttgatgag actctaagta cacttcagtt tgccagtact gccaaacatg tgagaaatac 1140
tccccatgtt aatgaggtcc tggatgatga agcgttgcta aaaagggtaca gaaaggaaat 1200
cttggattta aagaaacaat tagagaattt agagtcacgc tctgaaaca aagctcaagc 1260
aatggctaaa gaagagcata cacagttgct agctgaaatc aaacaactac acaaagagag 1320
agaagataga atatggcact tgacaaatat tgttgttgc tcatcccaag aatctcaaca 1380
ggaccaaagg gtcaaacgaa aacgaagagt tacgtgggag ccaggaaaaa tccaaaatag 1440
tttacatgct tctggtgttt ctgactttga tatgctatcc agattacctg gcaattttag 1500
caagaaggca aagttctctg acatgccttc atttccagaa attgatgact ctgtttgtac 1560
agagttttct gattttgatg acgcoctctc catgatggac agcaatggaa tagatgcaga 1620
atggaattta gccagtaaag taactcacag agaaaagaca tcacttcac aatcaatgat 1680
agactttgga cagattttctg acagtgttca gtttcatgat tcttctaagg aaaaccagct 1740
acaatacctc cccaaagact ctggtgatat ggctgaatgc agaaaagctt cttttgaaaa 1800
agagatcaca agcctccagc aacaactaca gtcaaaggag gaagaaaaaa aggaacttgt 1860
acaaagcttc gagctcaaga tagcagaact ggaagagcag cttagtgtca aagctaaaaa 1920
tctagagatg gttacaaact cgagagagca ttccattaat gctgaagtcc aaacagatgt 1980
tgaaaaggaa gttgtgagaa aagaaatgtc agtcccttga gactctggtt acaatgcatc 2040
aaacagtgac ctacaggata gttctgttga tggtaaactg ctaagcagct cccatgatga 2100
gtgtatagaa cacagaaaaa tgctggaaca aaagatcggt gatttagaag agtttattga 2160
aaaccttaac aagaaaagtg agaatagata acaaaaatct tctgagcaag attttatgga 2220
gagttatcag ctatgtgaag ctataatggc agaaaaggca aatgcactgg aggaactggc 2280
acttatcgca gataattttg acaatattat tctagagaat gaaactctaa aaagggaaat 2340
tgcagatctg gaacgttcac tcaaggagaa tcaagaaacc aatgagtttg aaattctgga 2400
gaaggaaact caaaaagaac acgaggcaca actaatccat gagattggca gtttaaagaa 2460
attagttgaa aatgcagaga tgtacaatca aaatcttgag gaagatctag aaactaaaac 2520
aaaacttctg aaagagcagg aaattcaact tgcagaatta aggaaacgcg cagataactt 2580
gcagaaaaaa gtacgaaatt ttgatctctc gggttccatg ggtgatagtg agaaactctg 2640

68

tgaagaaatc	tttcaactga	agcaatctct	ttctgatgct	gaagctgtga	ctcgcgatgc	2700
tcagaaggaa	tgttctttcc	tcagaagtga	aaatctagag	ctgaaggaga	aaatggagga	2760
cacatcaaac	tggtagaatc	aaaaagaaaa	ggctgcgtct	ttgtttgaga	agcagctgga	2820
aactgaaaaa	tcaaactaca	agaaaatgga	agctgatttg	cagaaagagt	tgcaaagtgc	2880
ttttaatgag	attaactact	taaatggcct	tctggcagga	aaggtcccca	gagatttgct	2940
ttctcgtggt	gaattagaga	aaaaggtttc	tgagttctca	aagcagcttg	agaaagcatt	3000
ggaagaaaaa	aatgccttgg	agaatgaagt	gacttgcccta	tcagaatata	aatttttgcc	3060
aaatgaagtt	gaatgcttga	aaaatcagat	cagcaaggct	tctgaagaga	taatgttatt	3120
aaagcaagaa	ggagaacatt	ctgcatctat	tataagcaaa	caagagatta	tcatgcagga	3180
gcaatctgag	cagattttac	aactgactga	cgaagtgaac	cacacacagt	caaaagtga	3240
gcagactgaa	gagcaatact	tggagatgaa	gaaaatgcat	gatgatcttt	ttgaaaagta	3300
tatcagaaac	aaaagtgaag	ctgaagacct	tttaagagaa	atggagaacc	ttaaaggcac	3360
tatggagtct	gtggaagtaa	agattgctga	cacaaaacat	gaacttgaag	aaactataag	3420
ggataaggag	caactgcttc	atgagaaaaa	atactttttt	caagcaatgc	agactatatt	3480
tccgattaca	cctctttcag	actcgccttc	tccctcaaaa	ttagttgaag	ggaactctca	3540
agaccccata	gaaatcaatg	actaccacaa	tttaatagcc	cttgctacag	aaaggaacaa	3600
cattatgggtg	tgtctagaga	ctgaaagaaa	cagtctcaag	gagcaagtta	ttgatttgaa	3660
cactcaactt	caaagtcttc	aagcacaaaag	tattgaaaag	tctgatctcc	agaagccaaa	3720
gcaagacttg	gaagaaggag	aggttaaact	gcttttgagg	atggaactac	tgaagggaca	3780
cctaactgac	tcacagctgt	ctatagagaa	attgcagctg	gagaatctgg	aagttacaga	3840
aaaactccaa	acacttcaag	aagagatgaa	aaatattact	atagaaagga	atgagcttca	3900
gaccaacttc	gaagacctga	aagcagagca	tgatagccta	aaacaagacc	ttagtgaaaa	3960
cattgagcag	tcaattgaaa	cacaagatga	attaagggtc	gccaggaag	agctaagaga	4020
acagaagcaa	ctggttgata	gcttttagaca	acagctttta	gattgttctg	taggaatttc	4080
atcaccaaac	catgatgcag	ttgctaacca	ggaaaagggtg	tcattgggtg	aagttaattc	4140
gttacaaagt	gaaatgctgc	gtggtgaaag	agatgagctt	caaacatctt	gtaaggcatt	4200
agtttcagaa	ctggagctac	ttcgtgctca	tgtaaaatct	gtggagggtg	aaaaccttga	4260
aattcacaaaa	aaactcaatg	gccttgaaaa	ggagatatatt	ggcaaactctg	aagaaagcga	4320
agtgttgaag	tccatgtttg	agaatctaaa	ggaagacaac	aataagctca	aagaacaagc	4380
agaggaatat	tctagttaaag	aaaatcaatt	cagtttagaa	gaggtgttca	gtggttcaca	4440
gaagctggta	gatgagatag	aggctcctgaa	agcacagcta	aaggcagcag	aagaaaggct	4500
ggaaataaag	gatagagatt	attttgaact	tgtacaaact	gcaaacacca	atthagttga	4560
gggaaaattg	gaaactccat	tgcaagctga	ccatgaggaa	gacagcattg	atcggcgttc	4620
tgaagaaatg	gagataaaag	ttcttgagga	aaaacttgag	cgaaatcagt	atctactgga	4680
aagattgcaa	gaagaaaagc	tggaactgtc	taacaaactt	gaaatccttc	agaaagagat	4740
ggagacgtcc	gttctattaa	aagatgacct	gcaacagaag	ctagaaagct	tgctgagtga	4800
aaacattatt	ctaaaagaga	atattgacac	aacctaaag	catcattcag	atactcaagc	4860
tcagctgcag	aaaacacagc	aagagctaca	gttggtctaag	aatcttgcaa	tcgctgcttc	4920
tgacaattgt	ccaataactc	aagaaaagga	aacctctgca	gattgtgtgc	atcctctgga	4980
agaaaagata	ttgttattaa	ctgaagaatt	gcatcaaaaa	actaatgaac	aggaaaaatt	5040
actacatgaa	aagaatgaac	ttgaacaagc	tcaggttgag	ctaaagtgcg	aggtggaaca	5100
tctgatgaag	agtatgatcg	aatcgaagtc	ctcacttgag	tccttgagc	atgagaaaca	5160
tgatactgaa	caacaacttc	ttgctcttaa	acagcagatg	caagtagtta	ctcaagaaaa	5220
gaaagagctg	caacaaaccc	atgaacactt	aacagctgag	gtggaccatc	taaaagagaa	5280
tattgaattg	ggtttgaatt	ttaaaaatga	agcgcagcaa	aagaccacta	aagagcaatg	5340
tctgctaaat	gagaataaag	aacttgagca	aagccagcac	agacttcaat	gtgagataga	5400
agagcttatg	aaaagcttaa	aggataaaga	gtcagcgctg	gaaactttta	aagaatctga	5460
acaaaaagta	attaactctga	accaagaaat	ggaaattggtt	atgctggaaa	tggaggaatt	5520
gaaaaaatagc	cataggactg	taattgctga	gagggaccag	ctgcaagacg	acctaaggga	5580
gagtgttgaa	atgtccattg	aaactcaaga	tgatctaaaga	aaggctcaag	aagcattgca	5640
gcagcagaaa	gataaagttc	aggaaactgac	ctcccagatt	tctgtcctgc	aggaaaagat	5700
ctctcttctg	gaaaatcaga	tgctttataa	tgttgcaact	gtgaaagaaa	ctctaagcga	5760
aagagatgac	ctgaaccagt	ctaagcaaca	cctgttctca	gaaattgaaa	ctcttagtct	5820
gtcttttaaa	gaaaagggaat	ttgcatttga	acaagcagag	aaggacaaag	ctgatgctgc	5880
caggaaaaca	atagatatca	cagagaagat	atcaaatata	gaagaacagt	tacttcaaca	5940
agccaccaat	ttaaaggaaa	ctctttatga	aagagagagt	cttatccagt	gtaaggagca	6000
actggctttg	aacacagaa	accttaggga	aacattgaag	agcaaagact	tggcattggg	6060
taaaatggag	caggagagag	atgaagctgc	caataaagta	atagctctta	cagaaaagat	6120
gtcttctcta	gaagaacaga	tcaatgaaaa	tgttactact	ctaaaagaag	gtgagggtga	6180
aaaagagacc	ttctaccttc	agagaccttc	aaagcagcag	tcttcttccc	agatggaaga	6240
gctcagagag	tcttttaaaga	ccaaagattt	gcagttggaa	gaggccgaga	aggagataag	6300

tgaagctact	aatgaaataa	agaatctcac	tgctaagatc	tcttctctag	aggaggagat	6360
tcttcagaat	gctagcattt	tgaatgaagc	tgtaagcgaa	agggaaaacc	ttcgccattc	6420
gaagcagcaa	ctgggtttcag	aattggagca	gctatcactg	acattaaaga	gtagagacca	6480
tgcattttgcg	caatctaaac	gcgaaaagga	tgaagctgta	aataaaatag	ccagtctcgc	6540
tgaagaaata	aagatcctga	caaaaagagat	ggatgaattc	agagattcaa	aggaatcctt	6600
gcaagaacag	tcttcccatc	taagtgaaga	gttatgtaca	tataagactg	aacttcaaat	6660
gctcaagcaa	cagaaagaag	acatcaacaa	caaacttgca	gagaaaagtt	aggaagtggg	6720
tgagctattg	caacacttat	catctctaaa	ggaacagctg	gaccaaatac	agatggagct	6780
aaggaatgaa	aagctcagaa	actatgaact	ctgcgaaaag	atggatatca	tggaaaaaga	6840
aatctcagtg	ctgcgtttta	tgcagaacga	gcctcagcag	gaagaagatg	atgttgacga	6900
acgtatggat	atacttgaga	gcagaaacca	agaaatacag	gagctgatgg	aaaaaatctc	6960
cgctgtgtat	tcagagcaac	acactttgct	cagcagctctc	tctagtgagc	ttcaaaagga	7020
aactgaagca	cacaaacatt	gcatgttaaa	tataaaggaa	tctctgtcat	ccacgctctc	7080
cagatccttt	ggcagcttgc	aaactgagca	tgtaagcta	aatactcaac	tgcagaccct	7140
tctgaacaaa	tttaaggttg	tataccgaac	tgctgcagtc	aaagaagatc	atagcttgat	7200
caaagattat	gagaaggacc	ttgctgctga	gcaaaagagg	catgatgagc	tgcgactcca	7260
actgcagtg	ttggagcagc	acggcagaaa	atggctcgat	tctgcatctg	aggaactcaa	7320
gttctgtgaa	attgaattct	tgaatgagtt	acttttttaa	aaagcaaata	taattcagag	7380
tgtccaggat	gacttttcag	aggtgcaggt	attcctaata	caagtaggat	caacactgca	7440
agaagagctt	gagcacaaga	aaggctttat	gcagtggttg	gaggaatttg	gagatctgca	7500
cgctgatgct	aagaaactca	gtgaaggcat	gcaacaggaa	aataggcgca	ttgcttctac	7560
catacagctc	ttaacaaaaa	ggctaaaggc	agttgttcag	tcgaaaatac	aacgtgagat	7620
aaccgtatat	ctgaaccagt	ttgaagcgaa	attgcaagag	aagaaagaac	aaaacaaaaga	7680
acttatgcgc	agaatggagc	accacggccc	tagtgctagt	gtaatggagg	aagaaaatgc	7740
tagactttta	ggcatactga	aaactgttca	agatgaatcc	aagaaactcc	aatcaaggat	7800
caaaatgcta	gaaaatgaac	tgaacttggt	caaagatgat	gccatgcaca	aaggtgaaaa	7860
agttgcaatt	ttgcaagaca	aactactaag	cagaaacgcc	gaagctgagc	taaacgcaat	7920
gcaggtgaaa	ctaactaaaa	agcaagataa	tcttcaggct	gcaatgaaag	aaatagaaaa	7980
cctacagaaa	atggttgcca	aaggtgcagt	accatataaa	gaagaaattg	acaaccttaa	8040
aactaagggtg	gtaaagattg	aaatggaaaa	aataaagtac	tcaaaagcaa	cagaccaaga	8100
gattgcctac	ttaaagctct	gtttggaaga	taaggaagaa	ggcttgcgta	ggttaaaaga	8160
ggaacttagg	cgagcacagg	cagacaacga	tacaacagtt	tgtgttccaa	aagattatca	8220
gaaagcttca	actttccctg	tgacttggtg	tggtggaagc	ggtatagtgc	agagcacagc	8280
aatgcttggtg	ctgcagtcgg	aaaaagccgc	cttggaaggg	gagctgtcac	attataagaa	8340
gaaatatcat	catttatcac	ggactatgtc	aagttctgaa	gatcgaaaga	aaacaaaggc	8400
aaaatctgat	gctcattctt	ctcatactgg	atcatcacac	agaggctcgc	ctcacaaaac	8460
tgaactttac	agacatggcc	ctgttactcc	agaaaggtct	gaaatgccaa	gcctacacct	8520
aggatctcca	aagaagtcgg	agtcacgac	taaacgtggt	gtgtcaccaa	acaggtccga	8580
aatctacagc	caattagtaa	tgtctccagg	caagaccggg	atgcataaac	atatactttc	8640
tccaagcaag	gttggaactgc	acaaaaagcg	tgccctgtct	ccaaacagat	cggagatgcc	8700
caccagcat	gtcatatccc	ctggcaagac	cggactgcac	aaaaatctaa	ctgaaagcac	8760
gttattcgac	aatttgtctt	ctccatgcaa	acagcaaaaa	gtacaggaaa	atctaaattc	8820
ccctaaaggc	aaattatttg	atgtgaagtc	aaaatcgatg	ccttactgtc	catctcagtt	8880
ttttgataat	tctaagcttg	gtgatttttc	agagctcaac	acagcagaga	gcaatgacaa	8940
aagtcaggct	gagaactggt	ggtatgaagc	aaaaaaagaa	acggcacctg	aatgtaaaac	9000
atcctagatc	cctgtacatc	tgactctcct	gtcctgcaaa	gagacttgct	actctgcctt	9060
cttgtaggaa	gaaacactag	aaactgccat	gtctgcataa	aggagtctca	ctggaagcaa	9120
aagttgttct	ttagtagtaa	tcactgggtg	ggcgagtggg	tacgtcttta	aaataaagtg	9180
caatacgtct	cacgtctatt	ttatatgtta	tgtctgtgta	tttgttacac	ttttaagtc	9240
cttgacttca	tatttggtct	atctgtagtt	tctttgtggt	tgcgcacaca	cacactggtg	9300
aaaaatgaca	tttgagctgt	attgttttac	tgactgggtc	ctctggggcc	atccatgcaa	9360
agcaccatta	gtgtgccaat	gtttttcact	acttattatt	atgtctgact	ttgtgaaata	9420
gaaaatacta	caaaagatag	ggcaaaagtt	gttttctcac	taggtaaacg	atatgggttt	9480
aacttaatta	ttctcctaata	aaaatattct	atataatggc	atctcattag	gtgacctaat	9540
ggttattctg	cactcttgta	aacctttttt	aaatttcact	cgtaataaag	cagccctgat	9600
tttaaatata	aaaaaaaaacg	gaattc				9626

<210> 3
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:9 amino acid
 insertion encoded by one cDNA isolated

<400> 3
 Asn Ser Arg Glu His Ser Ile Asn Ala
 1 5

<210> 4
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:putative
 nuclear localization signal (NLS)

<400> 4
 Arg Lys Lys Thr Lys
 1 5

<210> 5
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:primer

<400> 5
 gggctgccca ggaagag

17

<210> 6
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:primer

<400> 6
 gacagcattg atcggcg

17

<210> 7
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:primer

<400> 7
gaggggttcgg ccgctta

17

<210> 8
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 8
tctggggcca tccatgc

17

<210> 9
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 9
catatgacca tggccgaggg agatgcag

28

<210> 10
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 10
gtcaggtcag caacatacac g

21

<210> 11
<211> 30
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:sequence linked
at C-terminus of amino acids 1-473 of XCENP-E in
plasmid pET23dXCEMycHis

<220>
<221> PEPTIDE
<222> (12)..(20)
<223> c-myc epitope tag

<400> 11
Thr Val Ser Ile Ser Leu Gly Asp Leu Thr Met Glu Gln Lys Leu Ile
1 5 10 15

Ser Glu Glu Asp Leu Asn Phe Glu His His His His His
20 25 30

72